

Supplementary Information

Mutations in *TJP2* cause progressive cholestatic liver disease

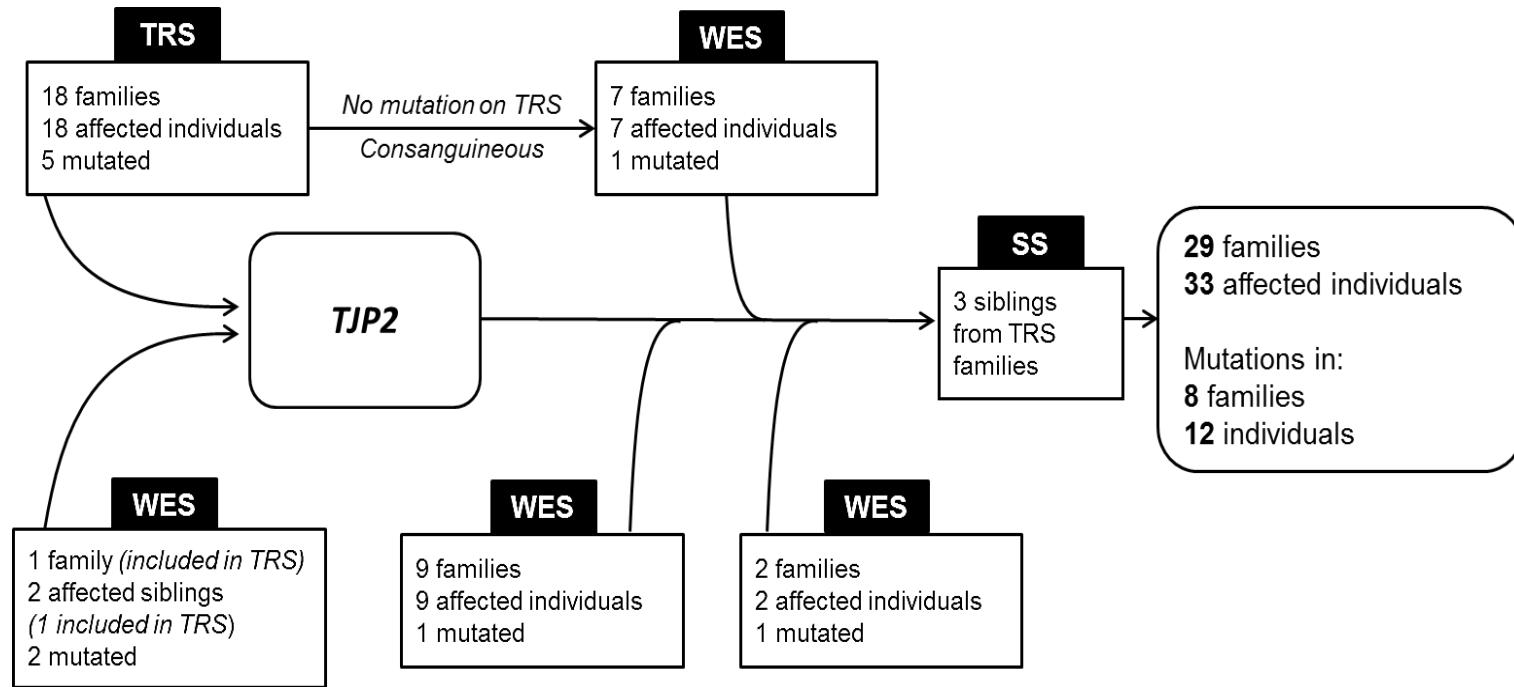
Melissa Sambrotta¹, Sandra Strautnieks², Efterpi Papouli³, Peter Rushton⁴, Barnaby E. Clark⁴, David A. Parry⁵, Clare V. Logan⁵, Lucy J. Newbury⁶, Binita M. Kamath^{7,8}, Simon Ling^{7,8}, Tassos Grammatikopoulos^{1,9}, Bart E. Wagner¹⁰, John C. Magee¹¹, Ronald J. Sokol¹², Giorgia Mieli-Vergani^{1,9}, University of Washington Center for Mendelian Genomics¹³, Joshua D. Smith¹⁴, Colin A. Johnson⁵, Patricia McClean¹⁵, Michael A. Simpson¹⁶, A.S. Knisely², Laura N. Bull^{17,18}, Richard J. Thompson^{1,9}

1.Institute of Liver Studies, Division of Transplantation Immunology and Mucosal Biology, King's College London School of Medicine, London, UK. 2.Institute of Liver Studies, King's College Hospital, London, UK. 3.Biomedical Research Center Genomics Core Facility, Guy's and St Thomas' National Health Service (NHS) Foundation Trust and King's College London, London, UK. 4.King's College Hospital NHS Foundation Trust, Department of Haematological Medicine, London, UK. 5.Leeds Institute of Molecular Medicine, St James's University Hospital, University of Leeds, Leeds, UK. 6.Renal Sciences, Division of Transplantation Immunology and Mucosal Biology, King's College London School of Medicine, London, UK. 7.Division of Gastroenterology, Hepatology and Nutrition, The Hospital for Sick Children, Toronto, Ontario, Canada. 8.Department of Pediatrics, University of Toronto, Toronto, Ontario, Canada. 9.Paediatric Liver, GI and Nutrition Centre, King's College Hospital, London, UK. 10.Histopathology Department, Royal Hallamshire Hospital, Sheffield, UK. 11.Department of Surgery, University of Michigan Medical School, Ann Arbor, Michigan, USA. 12. Department of Pediatrics, Children's Hospital Colorado and University of Colorado School of Medicine, Aurora, Colorado, USA. 13.Full lists of members and affiliations appear in the **Supplementary Note**. 14.Department of Genome Sciences, University of Washington, Seattle, Washington, USA. 15.Children's Liver and Gastroenterology Unit, Leeds General Infirmary, Leeds, UK. 16.Division of Genetics and Molecular Medicine, King's College London School of Medicine, Guy's Hospital, London, UK. 17.Liver Center Laboratory, Department of Medicine, University of California, San Francisco, San Francisco, California, USA. 18.Institute for Human Genetics, University of California, San Francisco, San Francisco, California, USA

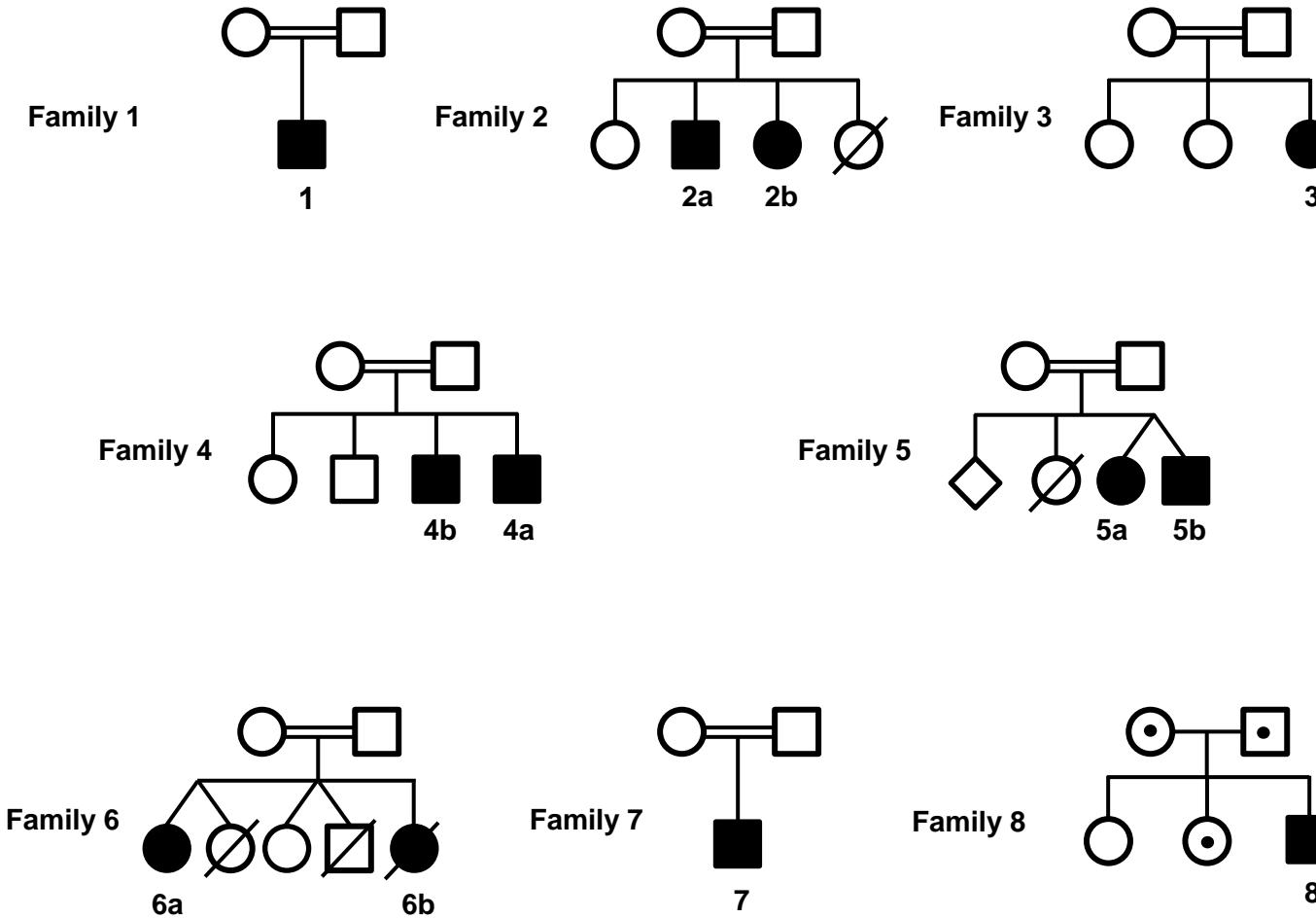
Supplementary Note	2
Supplementary Figure 1.	3
Supplementary Figure 2.	4
Supplementary Figure 3	5
Supplementary Figure 4.	6
Supplementary Table 1.....	7
Supplementary Table 2.....	9
Supplementary Table 3.....	10
Supplementary Table 4a.....	11
Supplementary Table 4b.....	12
Supplementary Table 4c	13
Supplementary Table 4d.....	14
Supplementary Table 4e.....	15
Supplementary Table 5.....	16
References	17

Supplementary Note

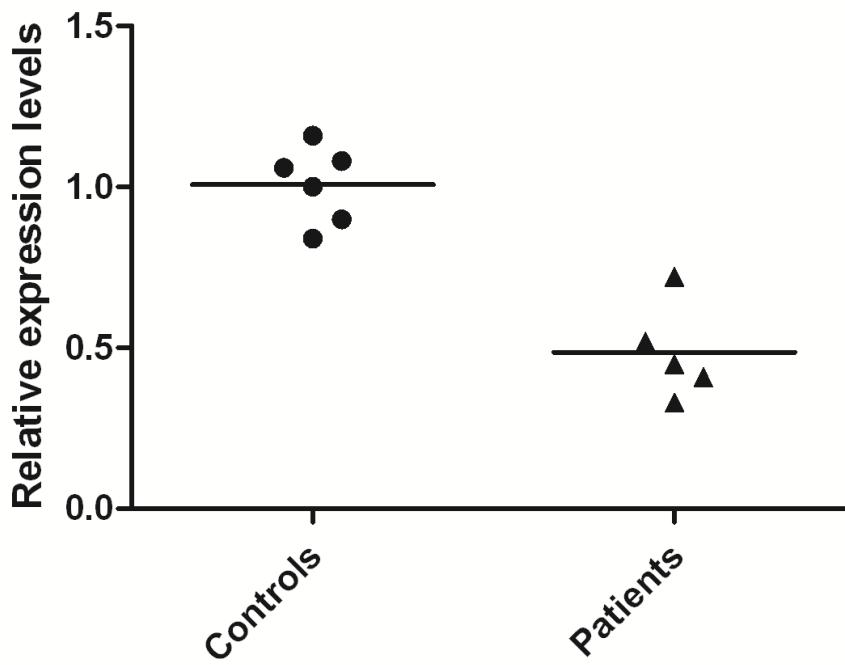
Whole-exome sequencing undertaken by the University of Washington Center for Mendelian Genomics (UW CMG) was directed by Drs. Debbie Nickerson, Jay Shendure, and Michael Bamshad.



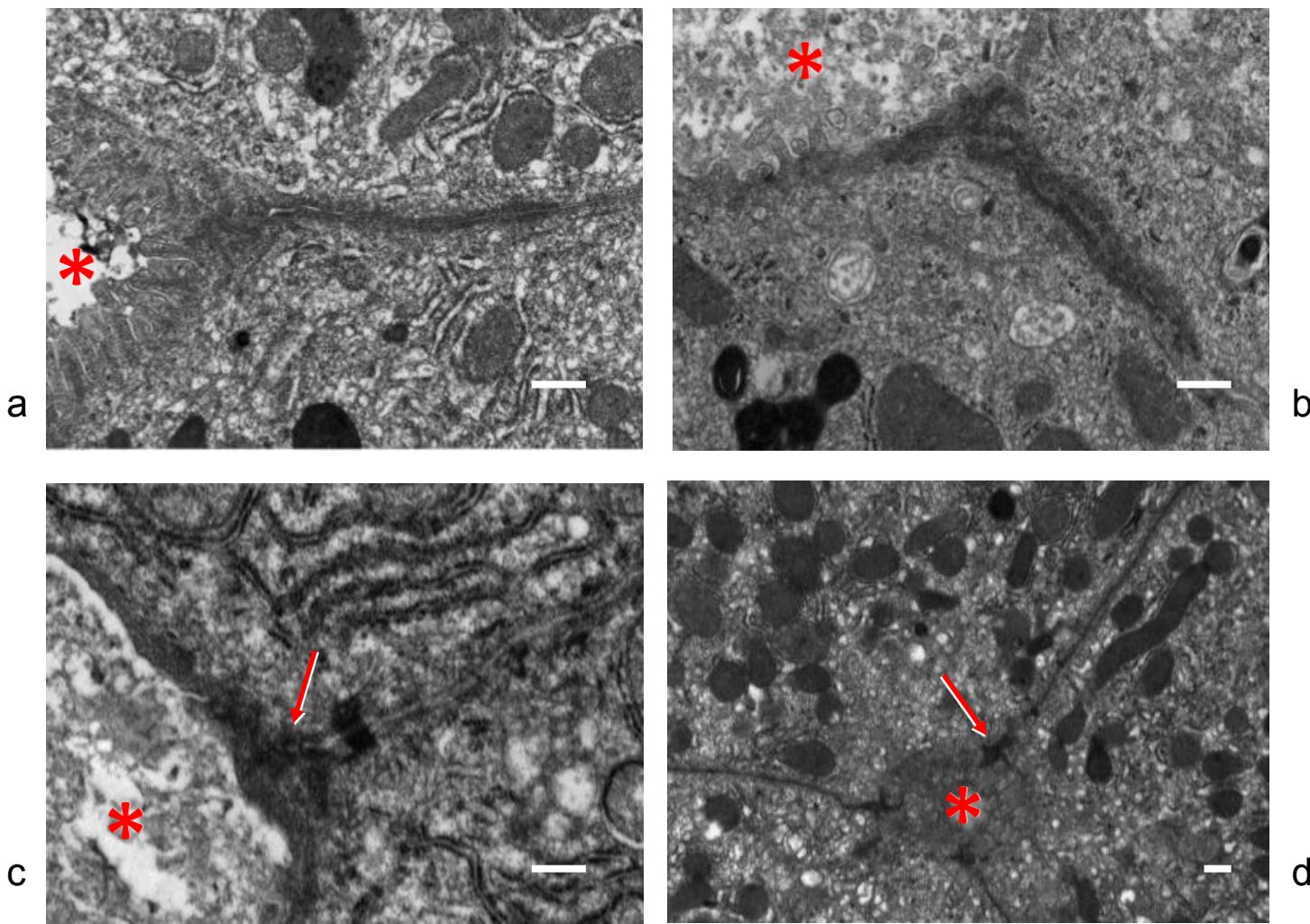
Supplementary Figure 1. Flow chart of work undertaken. Targeted resequencing (TRS), whole-exome sequencing (WES) and Sanger sequencing (SS) was conducted in one centre. Whole-exome sequencing was conducted in three other laboratories, as shown. Sample numbers for each are indicated. *TJP2* was identified as the disease-causing gene by target resequencing and whole-exome sequencing independently. All subsequent analyses were conducted in light of this knowledge.



Supplementary Figure 2. Pedigrees of the eight families found to be mutated in *TJP2*. Mutated individuals are indicated with filled shapes. Unfilled shapes are used for untested relatives. DNA was not available from unaffected individuals, with the exception of family 8 where 3 unaffected individuals were tested. All 3 were heterozygous for the disease causing mutation and are marked with dots. The deceased child in family 2 died of cardiac disease. The cause of death in family 5 is not known. The three dead children in family 6 both had chronic obstructive pulmonary disease and cholestatic liver disease.



Supplementary Figure 3. Quantitative analysis of *TJP2* mRNA levels in five patients and six controls. *TJP2* expression levels were measured in liver tissue by Taqman-based quantitative RT-PCR. The expression in all is expressed relative to the mean of the control samples. Each data point represents a single liver sample, though all were tested in triplicate.



Supplementary Figure 4. Transmission electron microscopy of tight junction structure in liver biopsies. All four panels show tight-junction complexes between adjacent hepatocytes and biliary canaliculi. In each panel a red asterisk indicates the canalicular space. Panel a is from a liver biopsy specimen obtained at presentation in patient 1. Panel b is taken from the explanted liver of patient 5b. Panel c is from a patient with BSEP deficiency and panel d from a patient with FIC1 deficiency. Tight junctions appear to extend deeper into the paracellular, or lateral, space in the *TJP2*-mutated patients, with diminution of the most electron-dense part of the zona occludens (shown with arrows in c and d). In all cases the scale bar = 500 nm; OsO₄/ uranyl acetate / lead citrate.

Gene	OMIM ID	Disease/Function	Reference
<i>ABCB11</i>	603201	PFIC	1
<i>ATP8B1</i>	602397	PFIC	2
<i>ABCB4</i>	171060	PFIC	3
<i>JAG1</i>	601920	Alagille syndrome	4,5
<i>NOTCH2</i>	600275	Alagille syndrome	6
<i>VPS33B</i>	608552	ARC	7
<i>VIPAS39</i>	613401	ARC	8
<i>TJP2</i>	607709	FHCA	9
<i>BAAT</i>	602938	FHCA	9
<i>ABCC2</i>	601107	Dubin-Johnson syndrome	10
<i>CLDN1</i>	603718	NISCH	11
<i>SERPINA1</i>	107400	A1AT deficiency	12
<i>CFTR</i>	602421	Cystic fibrosis	13
<i>SLC25A13</i>	603859	Citrin deficiency	14
<i>HSD3B7</i>	607764	HSD3B7 deficiency	15
<i>AKR1D1</i>	604741	AKR1D1 deficiency	16
<i>TMEM30A</i>	611028	FIC1 chaperone	17
<i>TMEM30B</i>	611029	FIC1 chaperone	17
<i>TMEM30C</i>	611030	CDC50 family member	18
<i>NR1H4</i>	603826	FXR	19
<i>ATP11C</i>	300516	Cholestasis (mouse model)	20

Supplementary Table 1. The 21 genes selected for NGS target resequencing associated with cholestatic liver disorders.

Abbreviations: PFIC, progressive familial intrahepatic cholestasis; ARC, arthrogryposis, renal dysfunction and cholestasis; FHCA, familial hypercholanemia; NISCH, neonatal ichthyosis / sclerosing cholangitis; A1AT, α -1-antitrypsin; HSD3B7, 3- β -hydroxy- δ -5-C27-steroid dehydrogenase; AKR1D1, δ (4)-3-oxosteroid 5- β -reductase; FIC1, familial intrahepatic cholestasis 1; CDC50, cell cycle control protein 50; FXR, farnesoid X receptor

	Patient	Gene Name	RefSeq ID	Nucleotide change	Predicted amino acid change	Zygosity	dbSNP
TJP2 deficiency patients	1	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Hom	rs2287622
		CFTR	NM_000492	c.91C>T	p.Arg31Cys	Het	rs1800073
	2b	ABCB11	NM_003742	c.1772A>G	p.Asn591Ser	Het	rs11568367
		ABCB11	NM_003742	c.1772A>G	p.Asn591Ser	Het	rs11568367
	3	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Het	rs2287622
		ABCB11	NM_003742	C.1213T>C	p.Cys405Arg	Het	-
	4a	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Hom	rs2287622
	5a	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Hom	rs2287622
	6a	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Het	rs2287622
		ABCB11	NM_003742	C.1213T>C	p.Cys405Arg	Het	-
Non-TJP2 deficiency patients	10	ABCC2	NM_000392	c.2535C>G	p.Tyr845Ter	Het	-
		ABCB11	NM_003742	c.1772A>G	p.Asn591Ser	Het	rs11568367
		CFTR	NM_000492	c.2758G>A	p.Val920Met	Het	-
	11	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Het	rs2287622
	12	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Hom	rs2287622
	14	HSD3B7	NM_001142778	c.71A>G	p.His24Arg	Het	rs201796375
		ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Het	rs2287622
	15	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Het	rs2287622
		TJP2	NM_004817	c.185C>T	p.Thr62Met	Het	rs138241615
	16	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Het	rs2287622
		CLDN1	NM_021101	c.404T>C	p.Val135Ala	Het	-
		TJP2	NM_004817	c.2810T>C	p.Leu937Pro	Het	rs28556975
	18	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Het	rs2287622
	19	ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Hom	rs2287622
	20	NR1H4	NM_005123	c.518T>C	p.Met183Thr	Het	rs61755050
		ABCB11	NM_003742	c.1331C>T	p.Val444Ala	Hom	rs2287622

Supplementary Table 2. Additional variants in *TJP2* and non-*TJP2* deficiency patients detected by targeted resequencing. Patients 1-6a and 9-20 were subject to TRS. Patients 7 and 8 are not included as they only had WES. The variants were selected having a minor allele frequency of less than 2% in the population in dbSNP 137, ESP, and 1000 Genome Project, and identified as pathologic using SIFT, PolyPhen and MutationTaster tools. In addition c.1331C>T in *ABCB11* has been included, as this common variant is known to reduce BSEP protein levels. Patients 9, 13 and 17 had no variants that met the inclusion criteria.

Abbreviations: Hom= homozygous; Het= heterozygous

	Patient 2a	Patients 2a and 2b	Patient 6a	Patient 7	Patient 8
Total variants in Exome Regions	178549	99854	54806	19027	49661
Filtered using dbSNP129 and dbSNP137 and ESP	13640	7539	4726	1435	2789
Not in 44 locally sequenced exomes	3680	1847	1761	608	1952
Non-synonymous, splice-site or indel	329	204	484	257	768
Damaging (condel, non-lof tolerant)	145	100	270	155	357
Homozygous	36	18	11	12	22*

Supplementary Table 3. Numbers of variants identified in patients 2a, 2b, 6a, 7 and 8 in whole-exome sequencing data.

Variants passing standard GATK hard-filters were removed in a series of steps:

- Variants present in dbSNP129 without “probable pathogenic” or “pathogenic” annotation in ClinVar, or in dbSNP137/NHLBI GO Exome Sequencing Project (ESP) with a minor allele frequency (MAF) equal to or greater than 2 %.
- Variants present in 44 locally sequenced exomes without PFIC
- Variants not resulting in non-synonymous, splice-site or exonic indels.
- Variants predicted as benign by condel²¹ or in loss-of-function tolerant genes²²
- Homozygous variants compatible with recessive inheritance.
- Genes frequently found to be mutated in local exomes (e.g. *MUC4*, *MUC6*, *DSPP*)

*Potential compound heterozygosity was considered for non-consanguineous family 8.

Gene	RefSeq ID	Nucleotide change	Predicted amino acid change	dbSNP
<i>FCRL1</i>	NM_052938	c.1138G>T	p.Val380Phe	rs148358454
<i>HMCN1</i>	NM_031935	c.4546C>T	p.Arg1516Trp	-
<i>TFB2M</i>	NM_022366	c.691G>A	p.Glu231Lys	-
<i>HAVCR1</i>	NM_012206	c.474_475insACA	p.Thr161dup	rs77147640
<i>CUL9</i>	NM_015089	c.1003C>T	p.Arg335Trp	-
<i>HOXA1</i>	NM_005522	c.212_223del	p.His71_His74del	-
<i>HOXA13</i>	NM_000522	c.496C>T	p.Pro166Ser	rs34398255
<i>UBAP2</i>	NM_018449	c.3068C>T	p.Thr1023Ile	-
<i>TJP2</i>	NM_004817	c.885delC	p.Ser296AlafsTer15	-
<i>ZNF503</i>	NM_032772	c.78ins12	p.Gly27_Alala28ins4Gly	rs72126859
<i>CYB5R2</i>	NM_016229	c.106C>T	p.Arg36Cys	-
<i>ATP12A</i>	NM_001676	c.3011G>C	p.Trp1004Ser	rs186599537
<i>IRF2BPL</i>	NM_024496	c.351insACA	p.Gln127dup	-
<i>PLCB2</i>	NM_004573	c.3061G>A	p.Ala1021Thr	-
<i>TARSL2</i>	NM_152334	c.587C>T	p.Thr196Met	rs148053074
<i>TTC19</i>	NM_017775	c.7C>T	p.Arg3Trp	-
<i>RAI1</i>	NM_030665	c.832_834delCGA	p.Gln291del	rs113303801
<i>HPS4</i>	NM_022081	c.266A>T	p.Asp89Val	-

Supplementary Table 4a. Candidate causal SNVs for family 2 identified in whole-exome sequencing data. Variants present in both affected individuals.

Variant list as summarized in Supplementary Table 3.

Gene	RefSeq ID	Nucleotide change	Predicted amino acid change	dbSNP
<i>CDC42EP3</i>	NM_006449	c.662A>T	p.Glu221Val	rs200264774
<i>PIKFYVE</i>	NM_001178000	c.373G>A	p.Ala125Thr	-
<i>ESPNL</i>	NM_194312	c.1474C>G	p.Gln492EGLu	rs140347701
<i>ESPNL</i>	NM_194312	c.2710G>A	p.Val904Met	rs200773816
<i>ACAD11</i>	NM_032169.4	c.1057C>T	p.Gln353Ter	-
<i>NOTCH4</i>	NM_004557.3	c.45_47del	p.Leu16del	rs35795312
<i>MAMDC2</i>	NM_153267.4	c.1085G>A	p.Arg362Gln	-
<i>IRF2BPL</i>	NM_024496.3	c.372_374delGCA	p.Gln127del	-
<i>RECQL5</i>	NM_004259.6	c.1586-2_1586-1insCA	p.?	rs142406301
<i>PLK5</i>	NM_001243079.1	c.569-1G>A	p.?	rs150328666
<i>MEX3D</i>	NM_203304.3	c.1531C>T	p.Pro511Ser	-

Supplementary Table 4b. Candidate causal SNVs for patient 6a identified in whole-exome sequencing data.

Variant list as summarized in Supplementary Table 3.

Gene	RefSeq ID	Nucleotide change	Predicted amino acid change	dbSNP
<i>IGFN1</i>	NM_001164586.1	c.5371G>A	p.Gly1791Arg	-
<i>FBLN2</i>	NM_001004019.1	c.1066G>A	p.Val356Met	rs200487515
<i>CYP21A2</i>	NM_000500.7	c.711T>A	p.Met240Lys	rs112398415
<i>USP20</i>	NM_006676.7	c.1075_1077delACG	p.Asp359del	-
<i>KRT1</i>	NM_006121.3	c.1669A>G	p.Ser557Gly	rs77846840
<i>GLIS2</i>	NM_032575.2	c.641G>A	p.Arg214Gln	-
<i>ATXN1L</i>	NM_001137675.3	c.566C>A	p.Pro189Gln	rs77031689
<i>FLCN</i>	NM_144606.5	c.1022C>T	p.Ser341Leu	-
<i>CILP2</i>	NM_153221.2	c.2380C>T	p.Pro794Ser	-
<i>GPI</i>	NM_001184722.1	c.913G>A	p.Glu305Lys	-
<i>KRTAP10-7</i>	NM_198689.2	c.135-2_135-1delAG	p.?	-
<i>SFI1</i>	NM_001007467.2	c.235C>T	p.Arg79Ter	rs202173110

Supplementary Table 4c. Candidate causal SNVs for patient 7 identified in whole-exome sequencing data.

Variant list as summarized in Supplementary Table 3.

Gene	RefSeq ID	Nucleotide change	Predicted amino acid change	Zygosity	dbSNP
<i>LCE4A</i>	NM_178356.2	c.143_144ins	p.Cys48_Gly49ins	Hom	-
<i>RHBG</i>	NM_020407.4	c.245T>G	p.L82R	Het	-
<i>RHBG</i>	NM_020407.4	c.1064del	p.Pro424GlnfsTer26	Het	-
<i>OBSCN</i>	NM_001271223.2	c.19949G>T	p.Arg6650Met	Het	rs80298121
<i>OBSCN</i>	NM_001271223.2	c.24121C>A	p.Pro8041Thr	Het	rs111553305
<i>OBSCN</i>	NM_001271223.2	c.25007C>T	p.Ala8336Val	Het	-
<i>KIAA1211L</i>	NM_207362.2	c.571G>A	p.Asp191Asn	Het	-
<i>KIAA1211L</i>	NM_207362.2	c.31C>T	p.L11Phe	Het	-
<i>MYO7B</i>	NM_001080527.1	c.3845G>A	p.Arg1282His	Het	rs201792059
<i>MYO7B</i>	NM_001080527.1	c.4844_4846del	p.Ser1615del	Het	-
<i>TJP2</i>	NM_004817.3	c.1894C>T	p.Arg632*	Hom	-
<i>PLCE1</i>	NM_001165979.1	c.67G>A	p.Val23Met	Het	rs144179807
<i>PLCE1</i>	NM_001165979.1	c.2671G>A	p.Gly891Ser	Het	rs199781223
<i>SLC22A18AS</i>	NM_007105.2	c.665C>T	p.Pro222Leu	Het	rs146094810
<i>SLC22A18AS</i>	NM_007105.2	c.599C>T	p.Thr200Ile	Het	rs139893801
<i>TROAP</i>	NM_005480.3	c.427C>T	p.Arg143Cys	Het	rs116282192
<i>TROAP</i>	NM_005480.3	c.1516C>T	p.Pro506Ser	Het	rs149122510
<i>ITGAE</i>	NM_002208.4	c.2815A>T	p.Ile939Phe	Hom	-
<i>ZZEF1</i>	NM_015113.3	c.724G>C	p.Glu242Gln	Hom	-
<i>C17orf49</i>	NM_001142798.2	c.331G>A	p.Asp111Asn	Hom	-
<i>SYNJ1</i>	NM_003895.3	c.806T>C	p.Phe269Ser	Het	-
<i>SYNJ1</i>	NM_003895.3	c.437G>A	p.Arg146Gln	Het	-

Supplementary Table 4d. Candidate causal SNVs for patient 8 identified in whole-exome sequencing data. Homozygous and compound heterozygous variations were considered in non-consanguineous family 8.

Variant list as summarized in Supplementary Table 3.

Abbreviations: Hom= homozygous; Het= heterozygous.

Patient	Chr	Start	Stop	Gene	No Exons	Read ratio	Type
2a	Chr 7	142478829	142561087	<i>EPHB6, PRSS2, TRY6</i>	18	0.0677	Del
6a	Chr 7	57530163	62758800	<i>LOC643955, ZNF716</i>	10	2.06	Dup
6a	Chr 9	71840241	71853695	TJP2	11	0	Del
7	Chr 1	152573269	152628815	<i>LCE3C, LCE3B, LCE3A</i>	3	0.0648	Del
7	Chr 4	75310853	75316250	<i>AREG</i>	4	1.88	Dup
7	Chr 4	8873247	9177894	<i>HMX1, LOC100288392, LOC650293</i>	5	5.55	Dup
7	Chr 7	62751898	62758771	<i>LOC643955</i>	7	2.12	Dup
7	Chr 7	74582344	74615899	<i>NCF1C, LOC100093631, GTF2IP1</i>	17	3.18	Dup
7	Chr 7	142481247	142482417	<i>PRSS2, PRSS3P2</i>	3	0.0693	Del
7	Chr 9	69403137	69447943	<i>ANKRD20A4</i>	13	0.0233	Del
7	Chr 9	71869127	71870121	TJP2	1	0.0117	Del
7	Chr 10	17840632	17898363	<i>TMEM236, MRC1 MIR511- 2, MIR511-1</i>	10	1.67	Dup
7	Chr 10	51843834	51845319	<i>FAM21B, FAM21A</i>	2	0	Del
7	Chr 15	45364517	45364637	<i>SORD, SORD</i>	1	0.12	Del
7	Chr 15	57730185	57754091	<i>CGNL1, CGNL1</i>	7	1.6	Dup
7	Chr 22	17175896	17265300	<i>XKR3</i>	11	1.89	Dup

Supplementary Table 4e. Candidate causal CNVs in patients 2a, 6a, 7, and 8 identified in whole-exome sequencing data. Analysis was performed using the ExomeDepth package. Potential CNVs with a Bayes Factor ≤ 10 were excluded.

Abbreviations: Chr= chromosome; Dup = duplication; Del = deletion.

a.

Patient No	Primer name		Primer sequence (5'-> 3')	Annealing temperature for PCR (°C)
1/2a/2b/3	Exon5	Fwd	GAACCTGAAACCTTAGTGAG	58
	Exon5	Rev	GTGTTGCTCTGTCATCCAG	
4a/4b	Exon9	Fwd	CTCTTGTACATGTCATTGTG	54
	Exon9	Rev	CAGCACATGTTCTGACTTT	
5a/5b	Exon14	Fwd	AGATTTCACAGAGACCCAG	54
	Exon14	Rev	GACTGTATGGTCAAGTTCTAAC	
6a/6b	Exon5	Fwd	GAGCATTGACCAGGACTA	56
	Exon17	Rev	TACATTTACATTTCACTTACCTG	
7	Exon 22	Fwd	GAAGCATCCTGATATCTATGC	58
	Exon 23	Rev	CCTTGCTTATGTCCTTCCAAT	
8	Exon 13	Fwd	TAGGACTTGTGCACTGAATG	56
	Exon 13	Rev	TATGTGAGAATGCGATCTCTG	

b.

Patient No	Primer name		Primer sequence (5'-> 3')	Annealing temperature for PCR (°C)
1	Exon 4	Fwd	CCCATGGAGGATGTGCTTC	56
	Exon6	Rev	AGGTTGCCATCTTAGTTGCC	
4a	Exon7	Fwd	GACTGTAACTGAGAACATGTC	56
	Exon11	Rev	GAATGCCAGCAACAAATATCC	
5a/5b	Exon12	Fwd	GAACACACAGGATTCAGAG	56
	Exon16	Rev	CTCCAGTGGATTCTCAGATC	
6a	Exon5	Fwd	CCTGGACCACGACTTTG	56
	Exon18	Rev	ATTTAGGTTGATTGTAGCTG	

Supplementary Table 5. Forward and reverse primers designed for *TJP2* (NM_004817).

(a) Intrinsic flanking primers used for Sanger sequencing with their respective annealing temperatures. (b) Exonic primers used for RT-PCR sequencing to investigate possible mRNA alternative splicing with their respective annealing temperatures

References

1. Strautnieks, S.S. *et al.* A gene encoding a liver-specific ABC transporter is mutated in progressive familial intrahepatic cholestasis. *Nat Genet* **20**, 233-8 (1998).
2. Bull, L.N. *et al.* A gene encoding a P-type ATPase mutated in two forms of hereditary cholestasis. *Nat Genet* **18**, 219-24 (1998).
3. de Vree, J.M. *et al.* Mutations in the MDR3 gene cause progressive familial intrahepatic cholestasis. *Proc Natl Acad Sci U S A* **95**, 282-7 (1998).
4. Li, L. *et al.* Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1. *Nat Genet* **16**, 243-51 (1997).
5. Oda, T. *et al.* Mutations in the human Jagged1 gene are responsible for Alagille syndrome. *Nat Genet* **16**, 235-42 (1997).
6. Kamath, B.M. *et al.* NOTCH2 mutations in Alagille syndrome. *J Med Genet* **49**, 138-44 (2012).
7. Gissen, P. *et al.* Mutations in VPS33B, encoding a regulator of SNARE-dependent membrane fusion, cause arthrogryposis-renal dysfunction-cholestasis (ARC) syndrome. *Nat Genet* **36**, 400-4 (2004).
8. Cullinane, A.R. *et al.* Mutations in VIPAR cause an arthrogryposis, renal dysfunction and cholestasis syndrome phenotype with defects in epithelial polarization. *Nat Genet* **42**, 303-12 (2010).
9. Carlton, V.E. *et al.* Complex inheritance of familial hypercholanemia with associated mutations in TJP2 and BAAT. *Nat Genet* **34**, 91-6 (2003).
10. Paulusma, C.C. *et al.* A mutation in the human canalicular multispecific organic anion transporter gene causes the Dubin-Johnson syndrome. *Hepatology* **25**, 1539-42 (1997).
11. Hadj-Rabia, S. *et al.* Claudin-1 gene mutations in neonatal sclerosing cholangitis associated with ichthyosis: a tight junction disease. *Gastroenterology* **127**, 1386-90 (2004).
12. Kueppers, F., Briscoe, W.A. & Bearn, A.G. Hereditary Deficiency of Serum Alpha-L-Antitrypsin. *Science* **146**, 1678-9 (1964).
13. Riordan, J.R. *et al.* Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* **245**, 1066-73 (1989).
14. Ohura, T. *et al.* Neonatal presentation of adult-onset type II citrullinemia. *Hum Genet* **108**, 87-90 (2001).
15. Cheng, J.B. *et al.* Molecular genetics of 3beta-hydroxy-Delta5-C27-steroid oxidoreductase deficiency in 16 patients with loss of bile acid synthesis and liver disease. *J Clin Endocrinol Metab* **88**, 1833-41 (2003).
16. Lemonde, H.A. *et al.* Mutations in SRD5B1 (AKR1D1), the gene encoding delta(4)-3-oxosteroid 5beta-reductase, in hepatitis and liver failure in infancy. *Gut* **52**, 1494-9 (2003).
17. Paulusma, C.C. *et al.* ATP8B1 requires an accessory protein for endoplasmic reticulum exit and plasma membrane lipid flippase activity. *Hepatology* **47**, 268-78 (2008).
18. Katoh, Y. & Katoh, M. Identification and characterization of CDC50A, CDC50B and CDC50C genes in silico. *Oncol Rep* **12**, 939-43 (2004).
19. Makishima, M. *et al.* Identification of a nuclear receptor for bile acids. *Science* **284**, 1362-5 (1999).
20. Siggs, O.M., Schnabl, B., Webb, B. & Beutler, B. X-linked cholestasis in mouse due to mutations of the P4-ATPase ATP11C. *Proc Natl Acad Sci U S A* **108**, 7890-5 (2011).
21. Gonzalez-Perez, A. & Lopez-Bigas, N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am J Hum Genet* **88**, 440-9 (2011).
22. MacArthur, D.G. *et al.* A systematic survey of loss-of-function variants in human protein-coding genes. *Science* **335**, 823-8 (2012).